The Bell System Technical Journal

Vol. XXVII *January, 1948* *No. 1*

An Experimental Multichannel Pulse Code Modulation System of Toll Quality

By L. A. MEACHAM and E. PETERSON

Pulse Code Modulation offers attractive possibilities for multiplex telephony via such media as the microwave radio relay. The various problems involved in its use have been explored in terms of a 96-channel system designed to meet the transmission requirements commonly imposed upon commercial toll circuits. Twenty-four of the 96 channels have been fully equipped in an experimental model of the system. Coding and decoding devices are described, along with other circuit details. The coder is based upon a new electron beam tube, and is characterized by speed and simplicity as well as accuracy of coding. These qualities are matched in the decoder, which employs pulse excitation of a simple reactive network.

I. Introduction

IN THE development of systems for transmitting telephonic speech, much effort has been directed toward minimizing the effects of noise picked up in the transmission medium. The system described in this paper represents one method which has been successful in eliminating completely such effects under appropriate and practical conditions. In this method, known as Pulse Code Modulation[1,2,3] (PCM), telephone waves are represented by sequences of on-off constant-amplitude pulses.

Perfect reception of such pulses demands simply recognition of whether any pulse exists or not. Recognition can be carried out effectively in the presence of noise and interference amounting to a substantial fraction of the pulse amplitude. In contrast, telephonic waves carry information by subtle amplitude variations in the course of time. High quality telephone reception accordingly demands a much lower ratio of noise and interference—lower by as much as 50 decibels.

The magnitude of this figure exhibits one good reason for exploring the possibilities of PCM. Another potent reason, which is of particular importance in multi-link transmission, is that, with pulses involving just two standard values, regeneration can be used at repeater points and at the receiver to wipe out transmission impairments. The regeneration process consists of the pro-

[1] A. H. Reeves, U. S. Patent 2,272,070.

[2] "Telephony by Pulse Code Modulation", W. M. Goodall, *Bell System Technical Journal*, July, 1947.

[3] "Pulse Code Modulation", H. S. Black and J. O. Edson; presented June 11, 1947 at the Montreal Summer Meeting of the American Institute of Electrical Engineers, and to be published in the *A. I. E. E. Transactions*.

duction of a properly formed standard pulse, free of noise, to correspond with each received pulse, even though the latter may be considerably misformed. The sole proviso here is that before regeneration the level of noise and distortion in each link be kept below the comparatively large threshold value at which a mark cannot be distinguished from a space. If this holds good throughout the transmission path then literally the received pulses can be made as good as new. In contrast it is impossible fully to repair or to regenerate signals not involving standard values of amplitude and of time. With such signals distortion and noise in each span contribute to the total which therefore increases with the system length.

To sum up, conversion of speech to a code of pulses and spaces permits telephony to assume certain new and desirable properties; ability to work with small signal-to-noise ratios, and ability to regenerate any number of times with no degradation of quality. These advantages do not accrue without certain penalties. Conversion of speech waves to pulse form and back to speech involves a certain degree of apparatus complexity at the terminals. This complexity is not decreased by the need to handle pulses at high speeds, of the order of a million per second. Here radar and television circuit techniques are helpful. Another characteristic is that a greater band width is occupied in the transmission medium. This arises through the operation of two factors, of which one is the use of double sideband in pulse transmission (as against single sideband in carrier telephony), and the second involves the number of pulses used in the code. The relatively wide band required can best be accommodated in the microwave region and it happens that the properties of on-off pulse transmission can be used there to particular advantage.

The PCM system to be described was set up to evaluate experimentally the problems involved in providing multichannel facilities of toll system quality. It was designed to accommodate 96 channels. For experimental studies of such things as crosstalk and methods of multiplexing channels, a fraction of the total number of channels is sufficient and only 24 of the 96 were built. These are arranged as two groups of 12 channels each. The channels of a group are assembled on a time division basis. Assembly of the groups is carried out on a frequency division basis, each group amplitude-modulating its own carrier. In a planned alternative arrangement the group pulses may be narrowed and interlaced to put all 96 channels in time division on a single carrier, but this alternative will not be explored here.

The assignment of 12 channels per group fits in well with the present arrangement of carrier telephone circuits used in the Bell System plant, such as Types *J*, *K*, and *L*.[4] Use of time division for a group of this size involves pulses with

[4] "A Twelve-Channel Carrier Telephone System for Open-Wire Lines," B. W. Kendall and H. A. Affel, *Bell System Technical Journal*, January, 1939. "Coaxial Systems in the United States," M. E. Strieby, *Signals*, January–February, 1947.

repetition rates up to 672 kilocycles, and pulse durations as short as a quarter microsecond. These pulses, obtainable from more or less standard types of vacuum tube circuits, can be distributed from point to point in the equipment without too much difficulty. Amplitude modulators and demodulators at two neighboring carriers—65 and 66.5 megacycles—then serve to bring the PCM signals into the intermediate frequency range for transmission to and from the microwave equipment.

The speech quality of the overall system in respect to such factors as band width, volume range, noise, distortion, and crosstalk more than meets the requirements generally imposed upon such systems.

Figure 1 is a front view photograph of the experimental apparatus setup with covers removed from one bay to show typical construction. The two end bays contain intermediate frequency modulators and demodulators required for the two groups. In addition voice frequency terminating sets and jacks are mounted here, together with testing equipment. The center bay and the one to the right of it are identical; each includes apparatus for handling a group of twelve message channels. Transmitting equipment is mounted in the upper half, and receiving equipment in the lower half of each bay. The remaining bay, second from the left, holds all the timing equipment needed to furnish control pulses for operating eight of the message bays, a total of 96 channels. Included are circuits for synchronizing the receiver. Individual regulated power supplies are mounted near their loads on the several bays.

Figure 2 is a rear view of the same equipment. Cables in the four horizontal ducts shown carry control pulses from the timing bay to the 12-channel group bays. These ducts are large enough in cross-section to handle all the cables required for a complete 96-channel terminal.

II. Functional Problems Involved

The broad problems brought together in building this system may be considered under the four classes following:

1. The pulse code modulation problem; to convert signal waves to pulse patterns.
2. The multiplex problem; to aggregate channels into groups and groups into a supergroup.
3. The transmission problem; to fit the system into the minimum required band width, and to remove the effects of transmission impairments.
4. The pulse code demodulation problem; to convert pulse patterns back to the original signal waves.

These are to be discussed from a functional standpoint to provide background for discussion of the specific equipment.

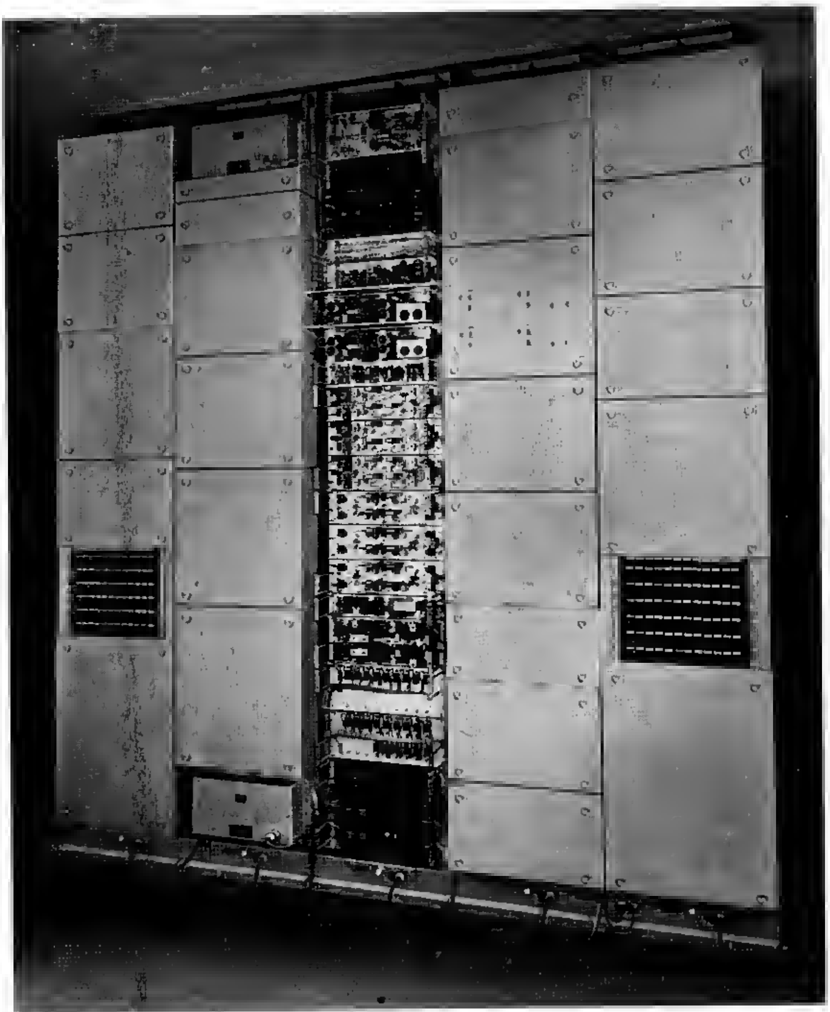


Fig. 1. Front view of experimental PCM terminal equipment, with covers removed from a 12-channel group bay.

Pulse Code Modulation

Sampling. A basic premise of pulse modulation systems is that the information content of a wave can be conveyed by samples taken at sufficiently frequent, equally spaced time intervals. The interval should be no greater than half the period of the highest frequency speech component to be reproduced or, otherwise expressed, the sampling rate should be not less than twice the frequency of the highest speech component present. This provision insures

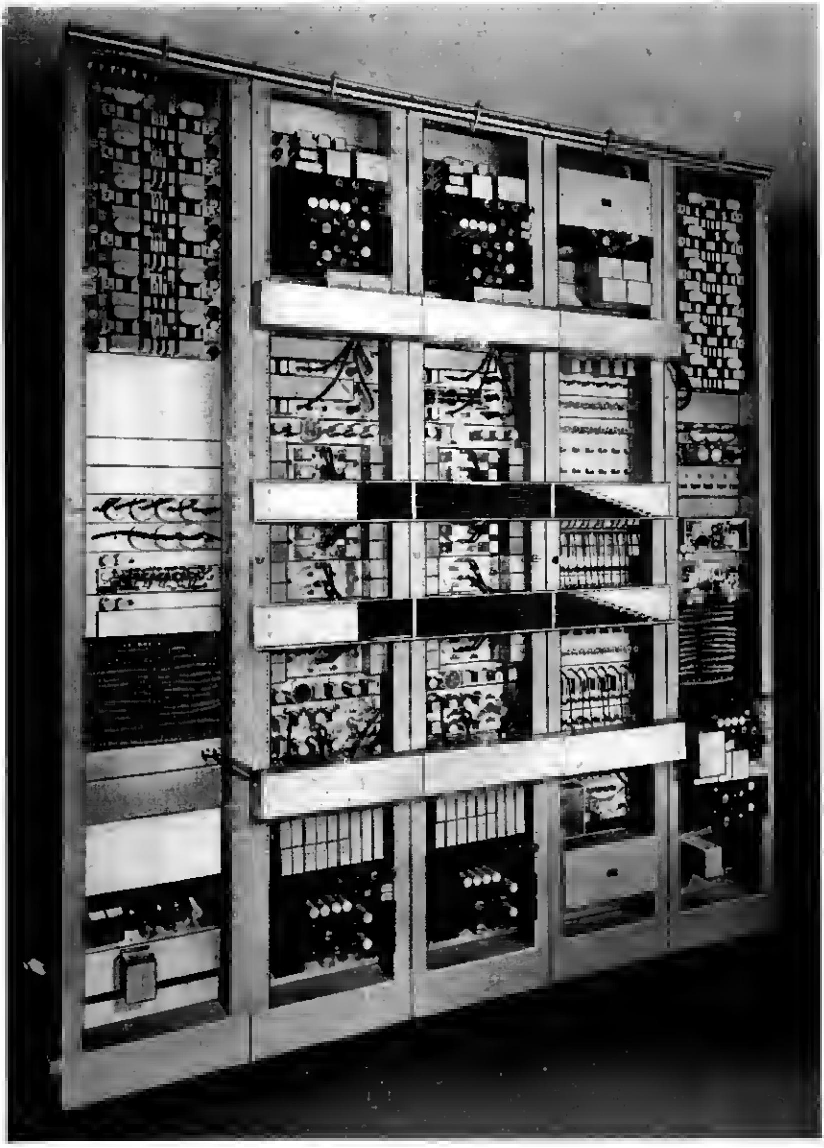


Fig. 2. Rear view. Cables in the horizontal ducts carry timing pulses leftward from the synchronizing bay.

that sidebands produced by sampling do not overlap to introduce distortion, as illustrated in Fig. 3. For a speech band extending up to 3400 cycles, a reasonable sampling rate is eight kilocycles.

Samples may be intermittently-transmitted portions of the signal wave, of appreciable duration, or they may be essentially instantaneous amplitudes.

To afford time for coding, the instantaneous samples may be maintained at constant value for an appropriate interval—a process here referred to as "holding."

Quantization. The fundamental operation of PCM is the conversion of a signal sample into a code combination of on-off pulses. In any practical system a continuous range of signal values cannot be reproduced since only a finite number of combinations can be made available. Each combination stands for a specific value, of course, so that we wind up by representing a continuous range of amplitudes by a finite number of discrete steps. This process is spoken of as quantization, a quantum being the difference between two adjacent discrete values. Graphically this means that a straight line representing the relation between input and output samples in a linear continuous

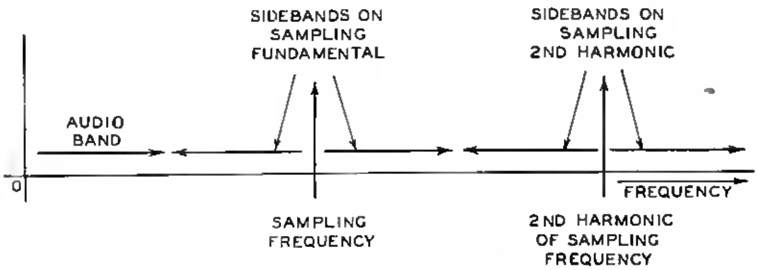


Fig. 3. Spectrum of a sampled audio band, illustrating separation of components when the sampling frequency is at least twice the top audio frequency.

system is here replaced by a flight of steps as in Fig. 4a. The midpoints of the treads fall on the straight line, and the height of the step is the quantum.

Manifestly the greatest error inherent in quantization amounts to half a step. Hence the quality of reproduction may be measured by the size of that interval, which depends upon the total number of steps in the amplitude range covered. With n pulses assigned to represent an amplitude range, the maximum number of discrete steps is 2^n, and the size of each step is proportional to 2^{-n} times the amplitude range.

This error shows up as a noiselike form of distortion, affecting background noise in the absence of speech, and accompanying speech as well. The distortion actually consists of a multiplicity of harmonics and high order modulation products between signal components and the sampling frequency scattered fairly evenly over the audio spectrum. If the audio signal is a simple sine wave, these many products may be identified individually; but for speech or other complex signals they merge into an essentially flat band of noise that sounds much like thermal noise. Since the level of this distortion is fixed by the quantum size, an adequate number of steps must be provided for the lowest

amplitude sounds it is necessary to transmit. Considering that consonant levels may be of the order of 30 decibels below vowels, and that weak talkers may be of the order of 30 decibels down from loud talkers, it is clear that amplitudes as far below maximum as 60 decibels require at least a few steps.

Ordinarily this would involve a large number of pulses for transmission, with a consequent increased complexity of terminal apparatus, and an in-

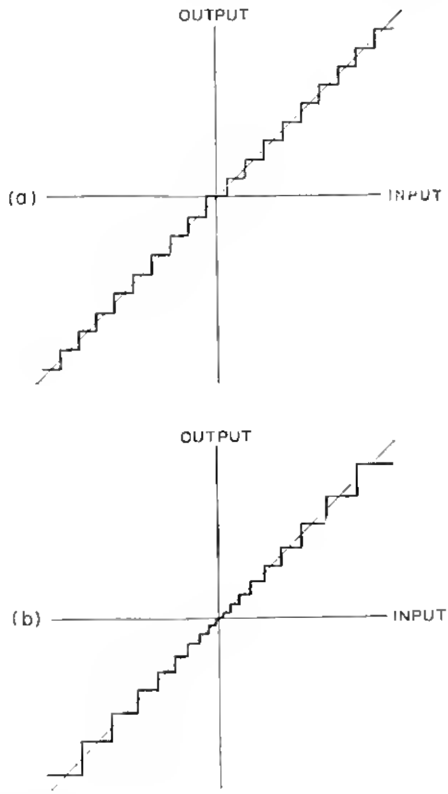


Fig. 4. Relation between input and quantized output, with quantization uniform in (a) and tapered in (b).

creased frequency band per channel. A more practical solution is to employ tapered steps rather than uniform ones. In this way a given number of steps can be assigned in greater proportion to the low amplitudes than to the highs, as shown in Fig. 4b. There results a degree of step subdivision sufficient to care adequately for the low-amplitude sounds including background noise. A small penalty is paid at the upper end of the amplitude scale because of the proportionately smaller number of steps available there, bringing in higher

quantizing noise in this range. How large the effect may be is determined by the degree of taper. To apportion the noise at different levels, the distribution of steps over the amplitude range can be varied.

Preliminary studies of quantization, involving listening tests and noise measurements for various numbers of digits and various kinds of taper, led to the choice of a seven-digit code (128 steps) for the present system. The taper employed reduces the smallest steps 26 decibels below the average size and increases the largest ones about 6 decibels.

Coding. The coder is required to set up a pulse code combination for each quantized signal value. A great many codes are conceivable, but in practice a simple one in which the pulses correspond to digits of the binary number system allows greatest simplicity at the receiver.

While coders may take a wide variety of forms, they can be arranged in three categories according to the way in which they evaluate speech amplitudes. In the first category an amplitude is measured by counting out, with a binary counter for example, the number of units contained in it one by one until the residue amounts to less than a unit. In the second, the amplitude is measured by comparison with one digit value after another, proceeding from the most significant digit to the least, and subtracting the digit value in question each time that value is found to be smaller than the amplitude (or its residue from the previous subtraction). In the third, amplitude is measured in toto by comparison with a set of scaled values. Modulators disclosed by Reeves[1] and by Black and Edson[3] are of the first category. That described by Goodall[2] belongs to the second, and the one described in the present paper is in the third category. Generally speaking the number of operations and the time required for coding decrease in going from the first to the third. Rapid coding is obviously desirable since it allows more channels to be handled in time division by common equipment.

Multiplex

Channels may be multiplexed by arranging them in time sequence, or by arranging them along the frequency scale. These methods are known as time division and as frequency division, respectively. The first is accomplished by gating, or switching, at precisely fixed times. One way of doing this impresses a more or less rectangular pulse on one of the grids of a gate tube, so that a signal wave may be transmitted during the gating pulse. The second method here refers to the use of amplitude modulators, each supplied with an appropriate carrier, which translate the signals to their assigned positions on the frequency scale. To avoid crosstalk in a time-division system, operations in group equipment common to a number of channels must proceed without memory of the amplitudes of preceding channels, requiring build up and decay times to be held within limits. This implies a sufficiently wide pass band

together with phase linearity. For the same purpose in a frequency-division system, filters are used to select and to combine channels. To avoid crosstalk the filters must be sufficiently selective, and amplitude non-linearity must be held within limits.

In the present system, the pulse code is delivered by the coder as on-off pulses in time sequence. It is therefore natural to organize the pulses of the different channels so that they appear in sequence, thus forming a time-division multiplex. Most types of coder require an appreciable length of time, after delivering the pulses of one channel, to prepare for coding the next. This preparation time may be afforded conveniently, without introducing gaps in the pulse train between assignments of consecutive channels, by providing two coders, which take turns at the channels in each time-division group.

As the number of channels increases, evidently the time interval which can be assigned to each channel must be reduced since all of them must be fitted into one period of an 8-kilocycle wave. Similarly the allowable duration of a code or digit pulse becomes shorter as the number of time-division channels in a group is increased. Then too, pulses tend to become more difficult to generate and transmit as their duration decreases. For these reasons it is desirable, and eventually it becomes necessary, to restrict the number of channels included within a time-division group. Frequency division may then be used to aggregate several time-division groups.

For our purposes groups of 12 channels are multiplexed by time division. With seven digits per channel, each group has a capacity of 672,000 pulses per second. To combine eight of the groups for a 96-channel system we again have a choice between frequency division and time division. The equipment of Fig. 1 is laid out to accommodate either procedure. In the first case each group is assigned a carrier for amplitude modulation, as used in actual tests to be described. For the second case the pulse durations would be cut down by a factor of eight, and the pulses from the different groups interlaced. Here the supergroup would have a capacity of 5,376,000 pulses per second.

In carrying out coding operations, and in multiplexing on a time-division basis, various control pulses are required which differ in repetition frequency, in time of occurrence, and in duration. These may be generated from a stable base frequency oscillator through the use of harmonics or, alternatively, of sub-harmonics. In a time-division system which requires a variety of flat-topped waves for switching operations, the use of sub-harmonics fits naturally. Frequency step-down circuits of the multivibrator type produce waves either approximating the desired forms directly, or requiring only simple circuits for reshaping. In contrast harmonic generation requires filters for component wave selection, more and more elaborate in structure as the order of the wanted harmonic goes up. Then, after selection, the harmonic has to be amplified and limited or otherwise shaped for switching purposes. Generally speaking

we need less apparatus and less space if we use multivibrator step-downs. Another advantage is that noise in the base frequency supply produces proportionately less phase jitter with sub-harmonics. While multivibrators do not have high inherent stability, they are capable of great precision when suitably controlled. For these reasons frequency step-downs are used to generate all the timing waves of the system.

Timing and gating operations require accurate time alignment of pulses, which may be accomplished by suitably delaying one set with respect to the other. For this purpose use is made of delay networks or cables, or delay multivibrators. Pulse durations may be varied through the use of interference effects between given pulses and their delayed replicas. Additional timing and gating wave forms are required for regeneration and for assembly in time-division multiplex. All such control pulses are economically generated at a single common point rather than by a number of local generators, and can then be supplied to the message equipment by common power amplifiers via shielded cable.

Transmission

In this section we are to consider the general factors entering into satisfactory reception of on-off pulses including such limitations as those on band width and noise. These are to be viewed while keeping in mind the procedures available for pulse regeneration.

If we start with a rectangular pulse like one of those generated at the transmitter, the corresponding frequency spectrum exhibits lobes extending indefinitely on the frequency scale, with progressively decreasing amplitudes as indicated in Fig. 5a. When such a pulse is passed through a linear phase filter[5] which discriminates against frequencies beyond the first lobe, the resulting pulse is practically of the sinusoidal form shown in Fig. 5b. Reducing the high-frequency response to the extent prescribed rounds the corners of the transmitted pulse, its duration at half value remaining equal to that of the original input pulse. In practice, transmission characteristics depart from phase linearity and low frequency cutoffs exist. Both effects introduce irregularities into the pulse response. While actual pulses therefore differ slightly in detail from the idealized picture given above, that picture will be retained for simplicity in discussion.

The band width needed for good pulse transmission can be minimized by making pulses as wide as possible. But since a specified number of pulses have to be put into a given time interval (84 pulses in $\frac{1}{8000}$ sec.) we must limit our broadening at a value where the presence or absence of a pulse may be clearly

[5] A suitable filter is one with a Gaussian characteristic, the loss in decibels varying as the square of frequency and having a value of 1 neper (8.68 decibels) at a frequency equal to 1/T.

determined in the presence of noise and interference. This spaces the sinusoidal pulses so that they overlap at their half-value points as illustrated in Fig. 5c. There it will be observed that no matter how many pulses occur in succession, the maximum amplitude of the pulse train is no different from that of a single pulse. The spectra of all such pulse trains have a common envelope, shown dashed in Fig. 5c.

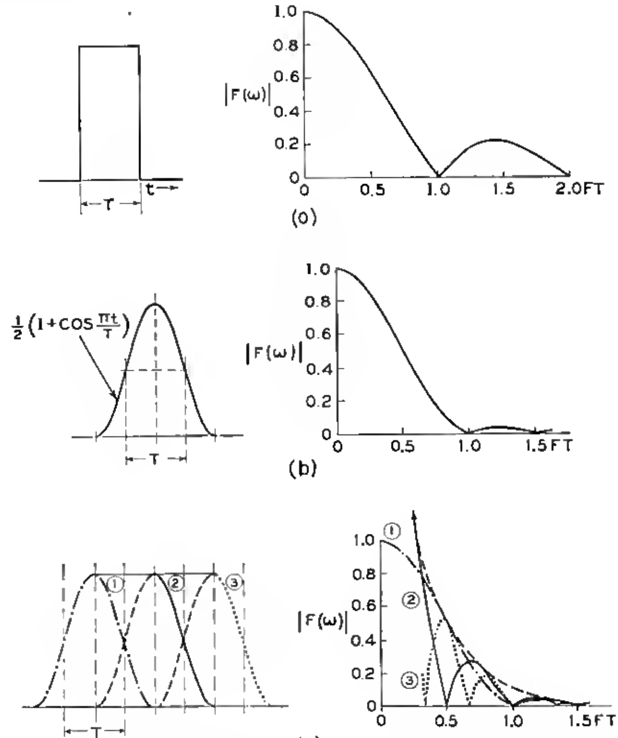


Fig. 5. Pulse forms and their associated amplitude spectra for: a, rectangular pulse; b, single lobe of a sinusoidally varying pulse; c, succession of pulses, each like that in (b).

Further, the band needed to transmit any sequence of such pulses is the same as that needed for a single pulse. This can readily be shown by using the familiar relation between transient build-up time and band width. Moreover, this conclusion is consistent with statistical analysis of all possible pulse combinations, which yields a spectrum of the form of Fig. 5b.

The relation between pulse duration and band width here described gives close to the optimum ratio of signal to noise and interference for the system considered. Narrowing the band would increase build-up and decay times, leading to reduced pulse amplitude and to increased pulse overlap. Thus,

although the narrower band would admit less extraneous noise, margins over noise and interference would be reduced. Widening the band, on the other hand, would allow pulses to build up and decay faster, but would not increase the pulse height. Thus the same signal would result, but the widened band would pass increased noise, and again the margins would be reduced. The optimum band width represents the most useful compromise in efforts to reduce noise and interference and to increase signal.

The filter characteristic we have been discussing is that of the entire link taking in all selectivity inserted between the practically rectangular pulses originally generated at the transmitter and the pulses delivered to the PCM receiver. Filters at both transmitting and receiving terminals of the link are included, the greater part of the overall selectivity being located at the receiver. With about 1.5 microsecond available per pulse at half amplitude, filters spaced 1.5 megacycle apart accommodate the double sideband and keep the interference between groups within tolerable limits.

To establish the presence of pulses we can set up an amplitude threshold equal to half the normal pulse amplitude, and test to see if that threshold is exceeded at a time near the center of the assigned pulse position. Selecting the threshold at half amplitude provides equal margin against the possibility of noise and interference bringing the full pulse amplitude, or mark, below threshold and bringing the nominal space above threshold. Testing at the pulse position midpoint maximizes this margin.

The amplitude threshold is set up by slicing a thin section horizontally out of the pulse at its half-amplitude level by means of an amplitude discriminator. Evidently, this procedure restores the flat top of the pulse. To complete regeneration by restoring the pulse epoch to a standard value the sliced pulses are gated at the midpoints of their proper intervals with narrow pulses supplied by the timing equipment. By these two pulse regenerating processes—slicing and gating—noise and interference are made impotent to produce errors until they attain a substantial fraction of the pulse amplitude. With the effects of noise and distortion thus eliminated, the only noise inherent in the system is that of quantization. In long systems having many repeater points, regeneration has to be practiced at spans short enough to permit cleaning out noise and distortion before it piles up above threshold. Thus in this system transmission impairments have to be considered only for the span between regeneration points; with their effects limited below threshold they are not carried over from one span to the next.

Where PCM groups are multiplexed by frequency division, amplitude nonlinearity of the system must be kept within limits. Otherwise intermodulation products may fall within transmission bands, adding to interference. Overlapping of neighboring filter bands also must be kept within bounds to reduce direct crosstalk between the pulse trains. With pulses arranged exclusively

in time-division multiplex, however, amplitude non-linearity of itself is not a factor. In this case the limitation comes on the departure from a suitable attenuation characteristic and from constant delay with respect to frequency. Distortion from these sources may increase the pulse duration so that excessive overlap of adjacent pulses will leave less margin available for interference and noise.

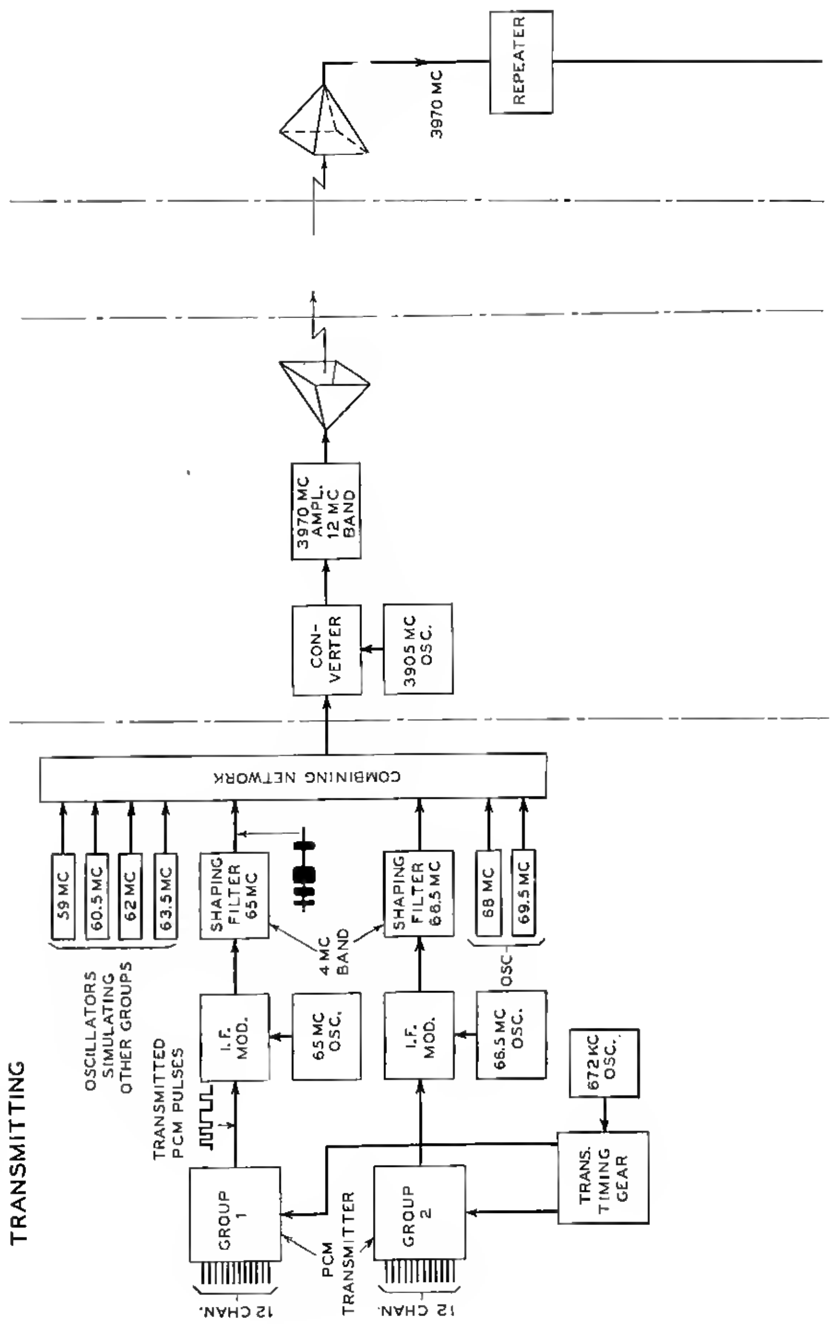
Pulse Code Demodulation

At the receiver, the regenerated code pulses are operated upon to recreate as closely as possible the original signal by operations complementary to those at the transmitter. Unfortunately, however, no process at the receiver can undo the effects of quantization which remain as noise, so that the quanta must obviously be made small enough from the beginning to permit satisfactory speech quality.

In each time-division group alternately working decoders may be used, in the same way as the two coders at the transmitter. Routing is effected by suitably timed gates. Conversion of a pulse code to amplitude may be accomplished by causing each pulse of a code combination to contribute an amplitude corresponding to the binary digit it represents, and then summing the contributions. When tapered steps are employed, due consideration must be given to overall linearity, discussed subsequently. The resulting output is a pulse-amplitude-modulated signal, which is then distributed to the channels of the group by an electronic commutator. Reconstruction of the signal from these distributed pulses is accomplished by simple filtering, which serves to remove components extraneous to speech introduced by the sampling procedure and tied up with the sampling rate.

Production of the necessary timing pulses at the receiving end proceeds in much the same manner as at the transmitter except for one thing. That is, instead of being initiated by a local oscillator, the receiver timing must be linked to the input pulses so that they may be properly routed. This involves the problem of synchronizing or, to borrow a term from television, framing. Use of this term is based upon the similarity of the sequence of PCM digits within a single sampling period to the complete ordered array of television picture elements. Preferably framing should be done with a minimum of time interval and of band width.

One method of synchronizing pulse systems employs a marker pulse which serves to initiate a timing sequence for each frame at the receiver. Here the marker pulse must differ sufficiently from the other pulses to permit its rapid and certain identification. This is ordinarily done by making the marker several times as long as any message pulse. In PCM, however, where digit pulses are run together in many code combinations, the marker would have to be very long to be clearly distinguishable, thereby cutting into channel ca-


TRANSMITTING
REPEATER
3970 MC
3970 MC AMPL. 12 MC BAND
CON-VERTER
3905 MC OSC.
COMBINING NETWORK
59 MC
60.5 MC
62 MC
63.5 MC
OSCILLATORS SIMULATING OTHER GROUPS
SHAPING FILTER 65 MC
SHAPING FILTER 66.5 MC
4 MC BAND
68 MC
69.5 MC
OSC
I.F. MOD.
65 MC OSC.
I.F. MOD.
66.5 MC OSC.
672 KC OSC.
TRANSMITTED PCM PULSES
TRANS. TIMING GEAR
GROUP 1
GROUP 2
PCM TRANSMITTER
12 CHAN.
12 CHAN.

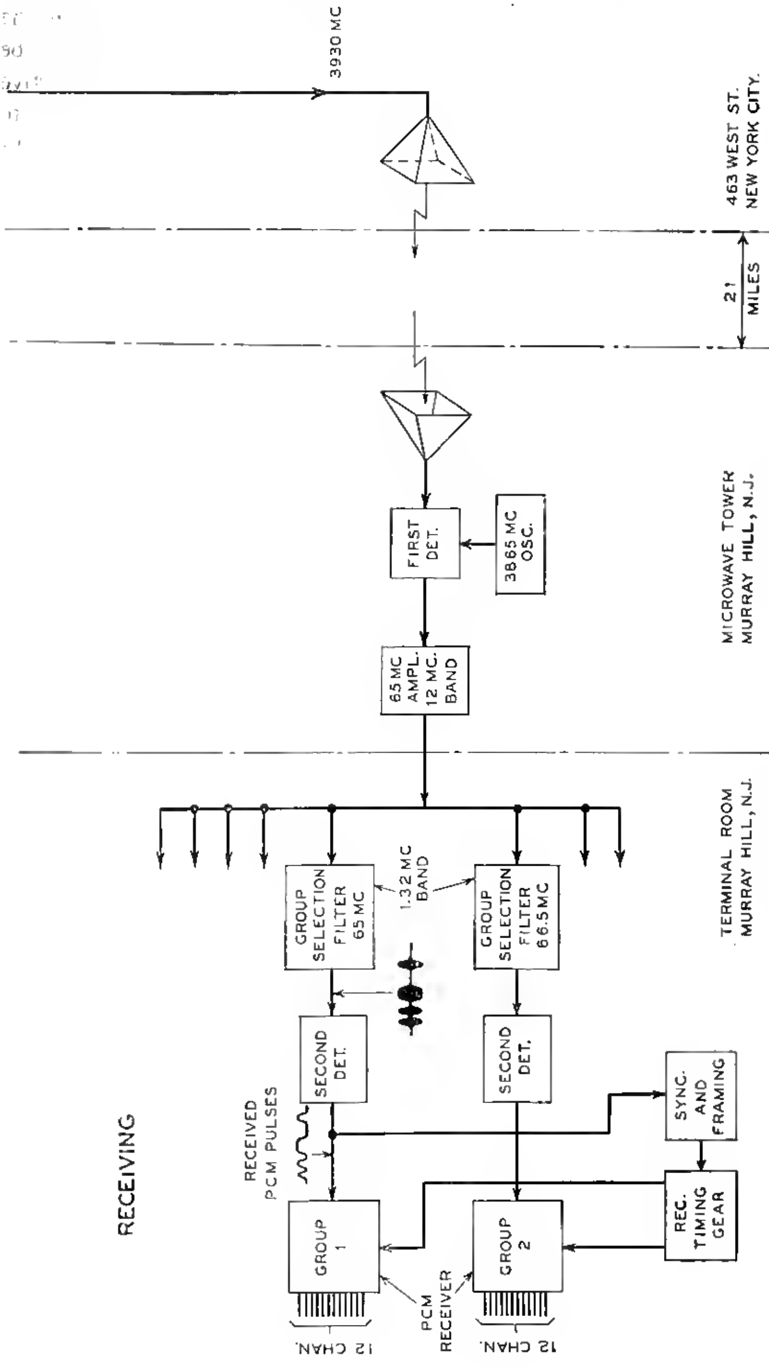


Fig. 6. Block diagram of the experimental PCM system.

pacity. This inefficiency can be avoided by deriving the pulsing rate from the pulse train through the use of a narrow band filter. Framing may then be effected by using in each frame just one digit pulse, which is given a distinctive repetition rate. With this method, a certain amount of time is required to establish synchronism when the system is started up. In the system to be described the framing time is less than one-tenth second—a tolerable value.

III. Experimental System

In the light of the foregoing discussion, the block diagram of the experimental system shown in Fig. 6 is believed to be largely self-explanatory. It will be noted that for microwave transmission the modulated intermediate-frequency signals are simply translated in frequency to the 4000-megacycle band. The shaping filters and group selection filters shown have approximately Gaussian characteristics, in accord with transmission considerations noted earlier. The band widths shown for these filters apply between points one neper down from the midband loss. Band widths given for the amplifiers, on the other hand, refer to their regions of essentially flat response.

Overall measurements are facilitated and the amount of experimental apparatus minimized by looping the radio path through a non-regenerative microwave repeater 21 miles away in New York. Both ends of the 24 channels are thus made available at Murray Hill, New Jersey, the location of the apparatus pictured in Fig. 1. With this arrangement, and using conventional 4-wire voice frequency terminating sets, 24 people are able to engage in 12 simultaneous conversations through the system.

PCM Transmitter. The transmitting equipment for an individual 12-channel group is shown schematically in Fig. 7. Each audio input[6] is passed through a 3400-cycle low-pass filter and through a limiter which chops off the positive and negative peaks of any signal exceeding a prescribed maximum amplitude. This limit is chosen to penalize the loudest talkers to the degree customary in toll system practice. The inputs then enter a "collector" circuit, which assembles samples of the channels in time division multiplex on a common lead. Although it functions electronically under control of pulses from the timing bay, the circuit so resembles a mechanical commutator that this analogy has been used in the schematic. The period of rotation of the "contact arm" is 125 microseconds (8 kilocycles), and a conducting path is formed to the common multiplex lead from each channel circuit in turn for $\frac{1}{12}$ of this period, or $10\frac{5}{12}$ microseconds. It should be

[6] In telephone terminology, these 4-wire inputs are normally at the −13 decibel level point; i.e., 13 decibels below the transmitting level at the toll test board. Strap connections are provided, as in the Western Electric A-2 channel bank, for adapting the system to inputs 3 decibels smaller. Similarly the final 4-wire outputs are delivered at the +4 (or +7) decibel point. The normal gain through a link is thus 17 decibels but can be set as much as 6 decibels greater.

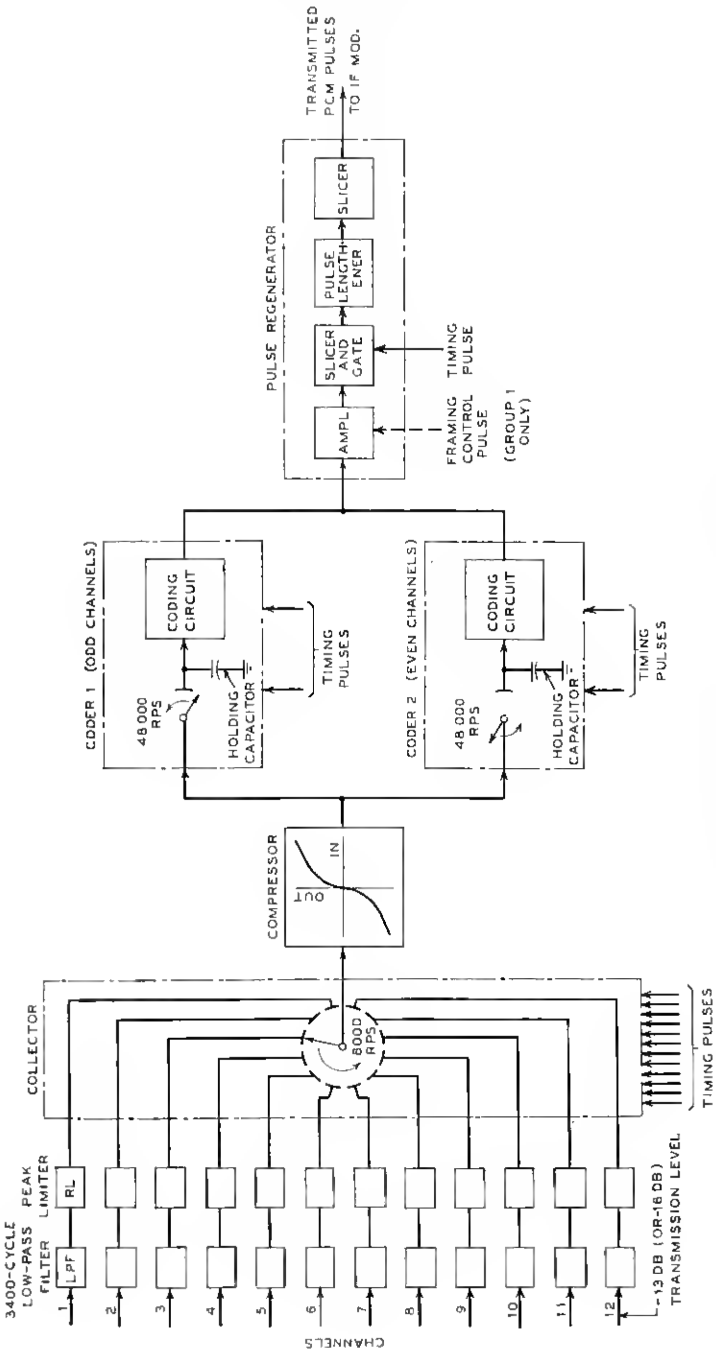


Fig. 7. PCM transmitter for a 12-channel group.

noted that the channel samples thus collected are not "held" at this stage; i.e., each sample does not remain constant in potential during its assigned interval, but rather changes to follow the wave form of the audio signal in the corresponding channel.

The multiplex signal is supplied to an instantaneous compressor, which employs silicon rectifier elements to give an input-output characteristic of the general form indicated within its block (Fig. 7). To understand the purpose of this device, we must recall the discussion of quantizing noise given earlier. There it was found desirable to provide a tapered distribution of step heights in the staircase-like quantizing characteristic, thus devoting a considerable number of small steps to the treatment of background noise and low-level signals. Although coders have been devised which inherently deal with signal amplitudes in this graded manner, it has heen found more practicable in the present system to apply amplitude compression to the samples before coding, and to divide this compressed amplitude range into uniform steps in the coder. The result is a tapered step distribution with respect to the original uncompressed scale of amplitudes, details of the distribution being determined by the shape of the compression characteristic.

It may be well to note here that this method can he used in reverse at the receiver, with the decoding performed on an equal-step basis, and the resultant samples passed through a complementary instantaneous expandor. If the compression and expansion are truly complementary, the overall characteristic relating amplitudes of input and output samples will he linear except for the tapered array of quantizing steps (Fig. 4b).

Incidentally, no added band width in the transmission path of this system is required to accommodate the instantaneous compandor action.

After compression, the multiplex signal is delivered at low impedance to the inputs of two coders. In Fig. 7 the switch analogy is called upon again to illustrate the routing of alternate samples to the "odd" and "even" coders, and concurrently the storage of these samples on "holding capacitors" to keep them unchanged during the coding operation. Here the switches rotate at 48,000 revolutions per second, each one closing six times in a complete 8-kilocycle frame. The contact segment is drawn as a short arc to indicate a brief closure, actually lasting ahout 5 microseconds and occurring while the switch of the collector is in contact with a single segment. When the circuit is thus completed from a particular channel to the holding capacitor, the voltage on the latter very rapidly assumes and then follows, for the remainder of the 5-microsecond interval, the potential of the compressed version of the channel signal. When the circuit opens, the latest state of charge, which is essentially an "instantaneous" sample, is left on the capacitor, and is thus held for ahout 16 microseconds—until the next closure. These sampling operations occur alternately in the two coders.

By a process to be considered later, each coder produces a 7-digit PCM code representation of its set of samples. The two coders deliver their code groups alternately on a common output lead during the final $10\frac{5}{12}$ microseconds of each 16-microsecond holding interval mentioned above. Individual pulses last about 1.5 microsecond, although as delivered from the coders they are somewhat irregular in timing and waveform. It will be noted that the interval allotted to the code group from each channel is just

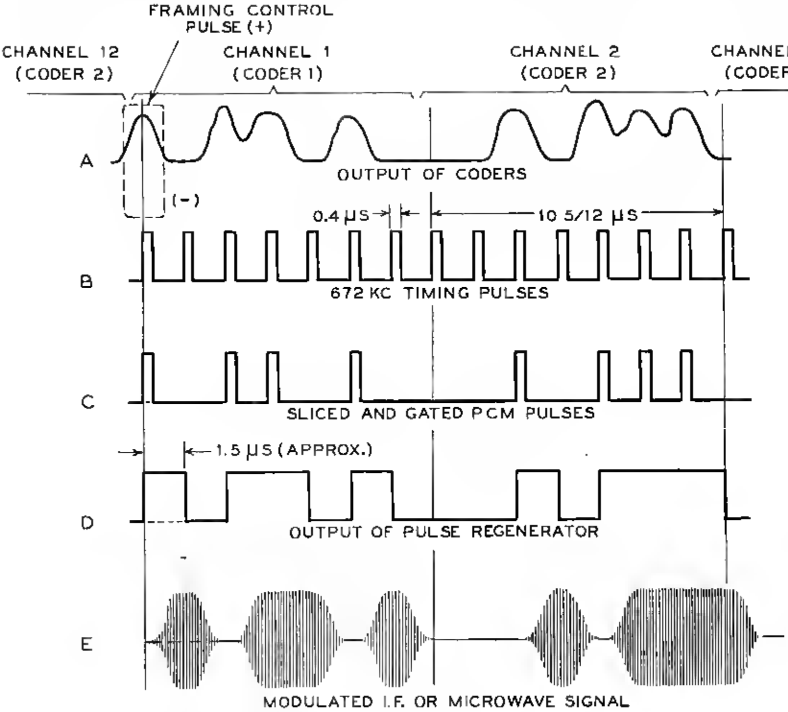


Fig. 8. Waveforms of the pulse regenerator.

$\frac{1}{12}$ of the 125-microsecond frame period and that a continuous train of code groups is thereby produced.

The common output circuit of the coders goes to a pulse regenerator which standardizes the pulses in height by slicing, and in time by gating, as illustrated in Fig. 8. The peaks of the coder output pulses (line A) are lined up in time with the gate control pulses (line B) supplied from the timing gear. The latter have a constant repetition frequency of 672 kilocycles and a uniform pulse length of 0.4 microsecond. Accordingly, the sliced and gated PCM pulses (line C) are also 0.4 microsecond long, and require lengthening to fill their allotted 1.5-microsecond intervals. This is accomplished by a

circuit in the pulse regenerator which first doubles the length of each pulse by adding thereto its own delayed reflection obtained from a short-circuited delay cable, and then slightly less than doubles it again by a similar process using a longer cable. A final slicing, to eliminate amplitude irregularities acquired in the lengthening process, yields square pulses as shown in line *D*, with adjacent pulses merged into a single longer pulse. This is the final output signal delivered to the intermediate-frequency modulator. Passage of the modulator output through a shaping filter results in rounded pulses (line *E*) suitable for transmission over the radio relay path.

In this regenerating apparatus, provision is also made for introducing a "framing control pulse," supplied from the timing bay and normally applied only to Group 1, although any other group may be used if desired. This pulse is about 1.5 microsecond long, and occurs once in each 8-kilocycle frame, but has opposite polarity in successive frames. It is timed to synchronize with the first digit of the Channel 1 code and is large enough in amplitude to override the pulse or space put out by the coder in that position. Hence in the final PCM output from Group 1, pulse 1 of Channel 1 is alternately present and absent regardless of the audio signal. This arrangement, used in automatic "framing" of the receiving timing gear as described in the following section, thus borrows the least significant digit from one channel of the system, leaving that channel usable, but with 6-digit instead of 7-digit quality. A 4-kilocycle tone of very low amplitude which it introduces in that single channel is made inaudible by the low-pass filter in the final audio output.

Synchronization. The connection of a transmitting channel to its proper receiving circuit in the time-division part of the system requires the two terminals to be synchronized: timing operations at the receiver must follow closely those at the transmitter. In a broad and general way this timing matter amounts to getting a local clock to keep the same time as a distant standard clock. Here the criterion of good timekeeping might be thought fussy by some standards; we cannot work with a discrepancy as long as a microsecond for the very good reason that incorrect routing of pulses would then result, associated with intolerably large decoding errors. Three provisions are made to take care of this situation. First, the framing is automatically monitored at all times. Second, if the system is out of frame—as it may be after transmission has been temporarily interrupted—the monitor circuit hunts for and establishes synchronism. Third, whenever the system is not properly synchronized and framed, all message circuits are cut off to avoid resulting noise and crosstalk.

For the purpose of this description we can use a mechanical analogy once more and picture all the transmitting channels of a time-division group arranged in order around a circle (Fig. 9). This time, however, we let each

small division of the circle constitute a commutator bar which is activated by the information of a particular code digit. Thus as the brush is stepped around at a uniform rate it puts digit information on the line in proper sequence. In each complete revolution there appears a set of on-off pulses constituting one frame. With twelve channels in the group, and seven digits to each channel, one revolution of the brush arm covers the frame of 84 digit positions. The revolution period is 125 microseconds. Roughly one and a half microsecond is allotted to each digit, and the brush steps 672,CC0 times a second. The driving force for stepping is supplied by a

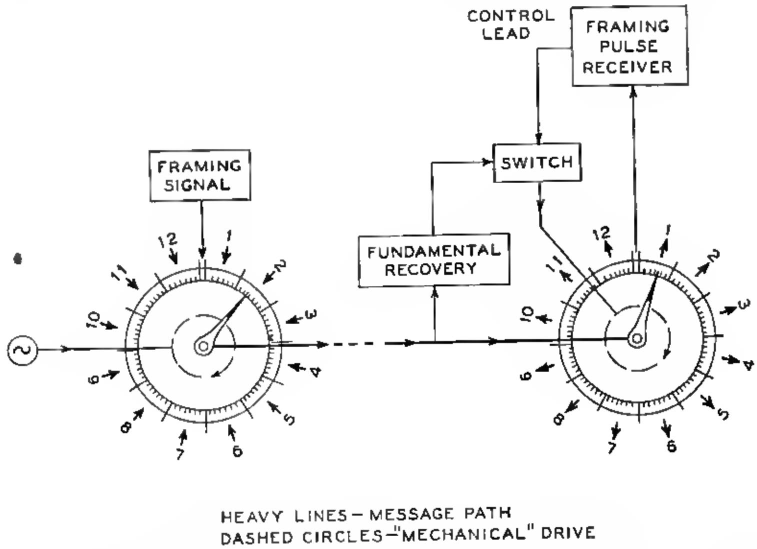


Fig. 9. Representation of the method used for framing, in relation to a single 12-channel time-division group.

stable oscillator, its frequency determined by a quartz crystal. One of the digit positions is taken away from its message duties and assigned for framing purposes, as described in the last section.

Our clock now is a one-handed affair with the brush arm for the hand. Its scale has the customary twelve main divisions, but each main division is divided into seven parts, one for each digit. Hour and minute markings refer to channel and digit, respectively.

At the receiving end we imagine a similar clock structure to be provided. Input pulses go directly to the clock hand which is stepped around to distribute pulses to the twelve message channels, plus a framing channel which takes up but a single digit space. Here the clock has no independent driving source, since it must follow the transmitter. Accordingly, the basic 672-

kilocycle frequency is recovered from the input pulse train by means which include filtering in a narrow-band crystal filter. The filter band is narrow enough to select the base frequency with negligible modulation and noise. That is, sidebands produced by pulse keying are attenuated to negligible proportions, as is the background noise. It is this output which serves to step the brush arm around the frame at precisely the transmitter frequency.

With the transmitter frequency recovered, both hands are stepped around their dials at the same rate. While this is a necessary condition for running the receiver, it is not a sufficient one since most likely the system will not be framed. In fact the odds are 83 to 1 against the two clocks indicating the same time if connection to the receiver is initiated at random times. If we had to deal with ordinary clocks, both in view, tbe resetting procedure could be accomplished by moving one clock-hand to agree with the other at one fell swoop. But in the PCM case resetting bas to be done more discretely since only one dial position per frame is viewed in the framing receiver. Accordingly, an orderly procedure is set up for locating the framing pulse which consists in examining digit positions one by one until the framing pulse is reached. After any one position is viewed long enough to establish the absence of the framing pulse, the receiving clock is set back one digit position and the next position viewed.

This resetting or framing procedure is governed by the framing receiver through its control of a switch which connects the recovered base frequency to the driving mechanism of the clock. If the channels are correctly routed, so that it is the framing pulse which is being viewed by the framing receiver, the switch is left closed, and the 672-kilocycle wave steps the clockhand around the dial without interruption. But if the system is not correctly framed the framing receiver does not get its distinctive pulses. In this case the switch is opened every little while for the duration of a single pulse interval, stopping the local receiver clock during that interval while the transmitter clock advances one digit position. In effect the receiving clock-hand is set back precisely one digit interval with respect to the standard. Thus the next digit pulse is brought into the framing receiver. If again the monitored input turns out to be other than the framing pulse, the stopping process is brought into play once more; this process is repeated until the system is framed.

As pointed out in the last section, the framing pulse is alternately absent and present in successive frames, corresponding to a 4-kilocycle rate. This is readily distinguishable from any of the message pulses, which in practice are found to have little energy content at this frequency. The framing receiver accordingly includes a resonant circuit tuned to four kilocycles. In the hunting process, eight frame periods are allowed between successive interruptions of the clock drive, to give tbe resonant circuit sufficient time

to build up above a threshold when the system is first framed. With this time interval thus fixed, each clock position is maintained for about one millisecond. Hence the time required to frame the receiver varies between one and eighty-three milliseconds, depending upon the epoch at which the system connection is established.

PCM Receiver. The received PCM signals of a 12-channel group are filtered from the frequency multiplex while in the intermediate-frequency state and then detected to the original signal band, as indicated in Fig. 6. They consist of rounded pulses, nominally sinusoidal in shape, but more or less distorted by transmission defects and accompanied by noise and interference.

These signals are supplied as input to the PCM receiver shown in Fig. 10. They are first sliced in amplitude, the slice being taken at approximately half the average or noise-free pulse height. Code groups of seven pulses are then routed alternately to two decoders, which handle even and odd channels respectively. The routing function is represented in the drawing by a two-segment commutator (A) rotating at 48,000 revolutions per second. Before entering the actual decoding circuit, the pulses are again sliced to secure very great uniformity, and are gated with 0.4-microsecond, 672-kilocycle pulses from the receiver timing equipment. Immediately after the arrival of the seventh digit of each code group of these standardized pulses, the decoding circuit produces a voltage on its low-impedance output lead proportional to the quantized amplitude represented by the code. As in the case of coding, details of the decoding process will be reserved for a later section of this paper.

The decoded amplitudes are available only momentarily; therefore it is desirable to sample each one at the proper time and store the result as a charge on a holding capacitor. This sampling process is represented by switch B in one decoder and B' in the other. Here the switch closures last only two microseconds, and values are held for about 19 microseconds.

The next step is to assemble the six samples from odd channels held successively by one capacitor and the six from even channels held by the other into a single time-division multiplex. Switch C performs this operation, rotating at 48,000 revolutions per second, and making contact alternately with the output circuits of the odd and even decoders.

This 12-channel multiplex signal is passed through an instantaneous expandor, the purpose of which has been noted in an earlier section. To simplify the problem of making the input-output characteristic of this circuit accurately complementary to that of the compressor, the two devices are designed to use identical silicon units. In the expandor, however, the non-linear device is employed in the feedback path of a broadband amplifier rather than in the direct transmission path, thus giving the inverse character-

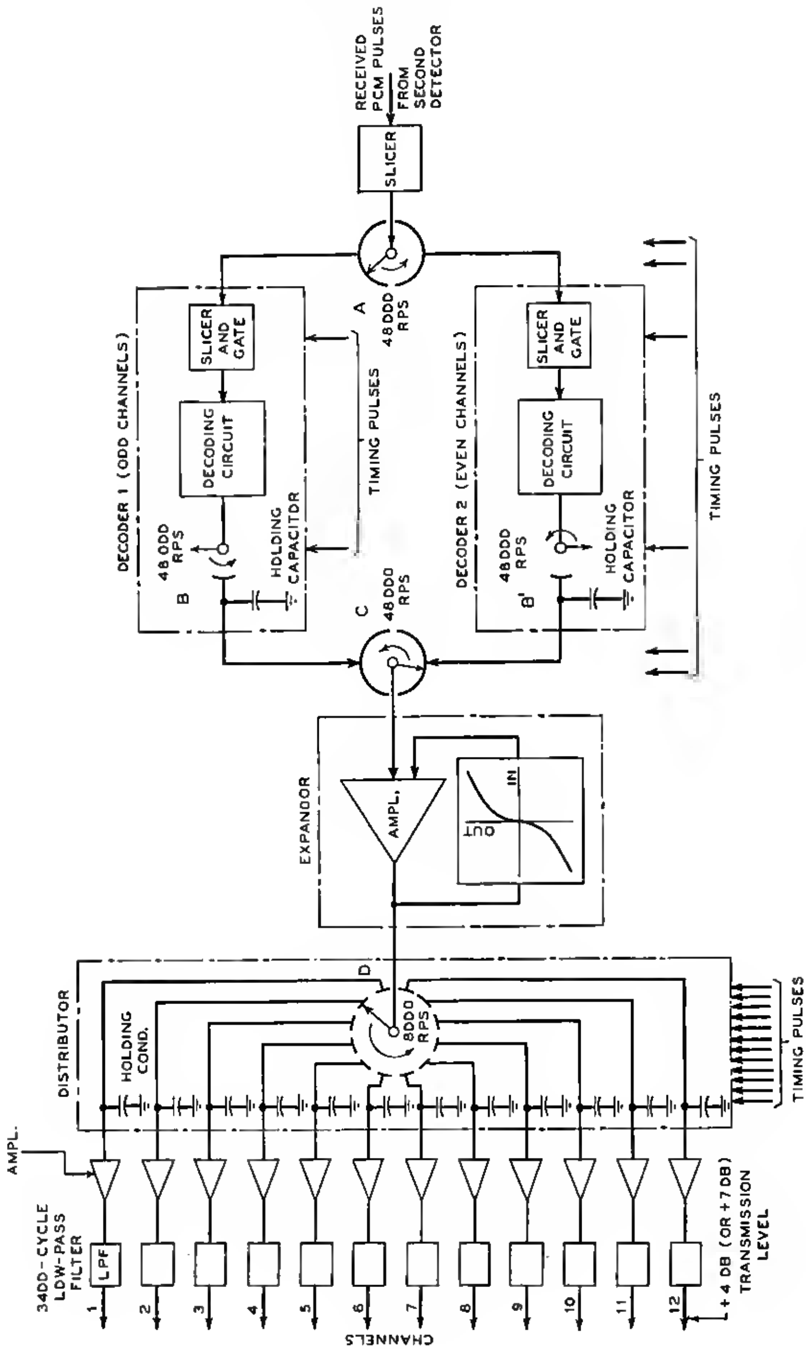


Fig. 10. PCM receiver for a 12-channel group.

istic. To allow any compressor to be used with any expandor, all the silicon elements are matched to a chosen standard unit, using selection and resistance "padding" methods. By such means, and by use of sufficient loop gain in the feedback amplifier, very satisfactory overall linearity of the system has been attained.

At the output of the expandor the waveform of the multiplex signal is essentially the same as that at the input of the compressor in the transmitting terminal. The samples are distributed to their respective channel destinations by a distributor *D*, resembling the collector described earlier. The rotation rate is 8000 revolutions per second, but the duration of contact on any one segment is only five microseconds instead of the possible full twelfth of the 125-microsecond frame period. This effective narrowing of the contact segments is done to allow the closure to occur well within the interval in which the circuit is completed by switch *C* from the output of a particular decoder. Each of the 12 segments of the distributor is provided with a holding capacitor, which stores its allotted samples for the full 125-microsecond frame period. The potential on any one of these capacitors thus changes at 125-microsecond intervals from one quantized sample amplitude to the next derived from the same original speech wave.

This potential is of sufficient magnitude to require use of only a simple single-stage triode amplifier for the output of each channel. Lengthening the samples by holding, as described, helps to make this possible by causing the amplifier to deliver useful power continuously instead of on a fractional time basis.

The only disadvantage of using lengthened pulses arises from an effect, very similar to the "aperture effect" encountered in sound movies, which introduces a curving slope across the audio gain-frequency characteristic of the system. In the present case the gain drops about three decibels as the frequency goes from the lowest to the highest value of interest. This slope can be corrected by a simple equalizing network, as shown in Fig. 11. In the present system the equalization is incorporated in the low-pass filter at the input of each audio channel. This is preferable to equalizing at the output, where power is at a premium.

The outputs from the channel amplifiers are passed through 3400-cycle low-pass filters, identical with the input filters except for omission of the equalization, and are delivered to standard voice-frequency circuits at the same levels[6] as are provided by a type *J*, *K*, or *L* carrier system.

IV. Component Circuits

Many of the circuit techniques used in the experimental system are conventional, others are more or less unfamiliar, and still others are believed to be novel. In the following some of the more important building blocks are

described. These include sampling circuits, the instantaneous compressor, slicers, and the PCM coder and decoder.

Sampling Circuits. The function of opening and closing a circuit at prescribed instants, represented by rotating switches in the block schematics, is actually performed[7] by one or the other of the devices shown in Fig. 12, employing diodes and triodes, respectively.

The diode type (Circuit a) is normally biased "open" by rectified charges stored on the two large capacitors. While in this condition it presents an extremely high series resistance between the low-impedance input and the load. But when a flat-topped pulse is impressed upon the pulse transformer, the aforesaid high impedance changes to a low value (of the order of 100

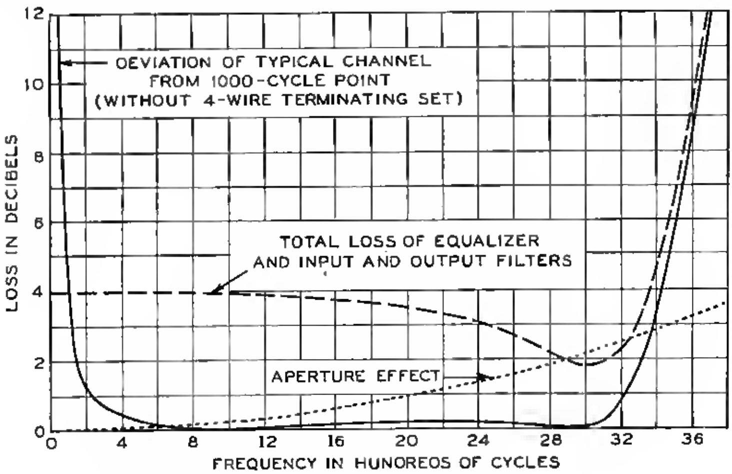


Fig. 11. Equalization for the aperture effect.

ohms), for the duration of the pulse. The upward-sloping tops of the pulses sketched in Fig. 12 represent predistortion to allow for transmission through the pulse transformer, for the purpose of obtaining rectangular pulses at the tube itself.

In the other variety (Circuit b) the plate-cathode paths of the two triodes are connected directly in parallel, conducting in opposite directions. The grids are both arranged to be biased below cut-off during the "open" condition of the sampler by grid rectification of pulses, and to be driven strongly positive by a pulse when a low-resistance conducting path is required between source and load.

These two types have their respective advantages. For example, the low capacitance to ground which the diode type affords across its load makes it

[7] A single exception is the case of switch A of Fig. 10. In this case two gated slicer circuits are used, as described subsequently.

preferable for use in the collector, where twelve samplers are multipled to a common load. On the other hand the triode type affords a d-c. path (without the series blocking capacitors that are required by circuit a) between source and load. In cases where a holding capacitor is to be charged to a succession of sample amplitudes which must be kept mutually independent to avoid crosstalk, the d-c. path is very desirable for it avoids "memory" effects associated with passage of the charging currents through the blocking capacitors. A further useful property of the triode circuit is its ability to

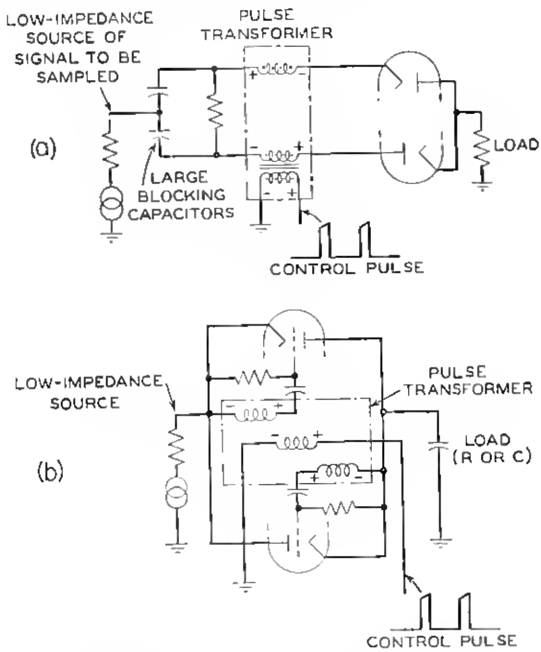


Fig. 12. Diode and triode sampling circuits.

sample signals of greater amplitude than that of the control pulses. With the diode circuit, the signal amplitude must be kept less than half the pulse height.

Instantaneous Compressor. The simple configuration of the non-linear circuit used in the compressor and in the feedback path of the expandor is shown in Fig. 13. Two selected silicon rectifiers are connected in parallel, but poled oppositely, with a small padding resistor (R_1 or R_2) in series with each. A parallel padding resistor (R_3) of large value is also provided. The direct-current resistance of this combination varies from about 6000 ohms at zero signal to about 190 ohms at the peak of a full-load signal. Input is applied as a current through the relatively high resistor R_4, and the voltage

drop across the varistor unit constitutes the compressed output. Although the development of this compressor has been a problem of many interesting aspects, it must suffice here to point out that copper-oxide elements are unsuitable at the speeds involved because of excessive capacitance, that silicon is superior to at least the presently available types of germanium

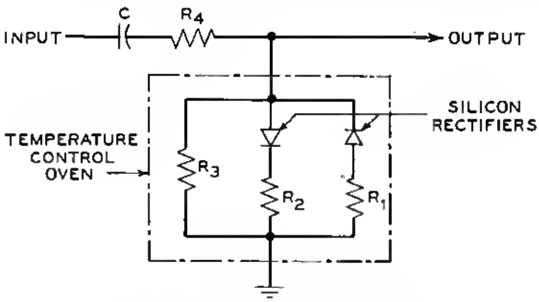


Fig. 13. Instantaneous compressor.

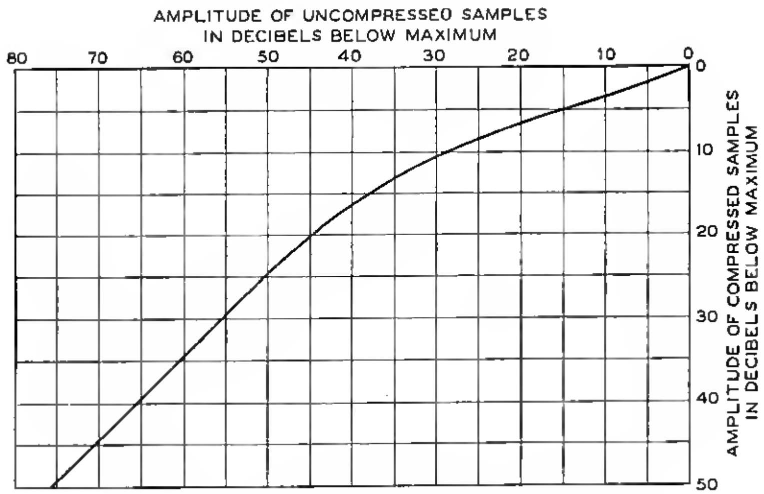


Fig. 14. Characteristic of the instantaneous compressor.

varistor in regard to freedom from certain memory or hysteresis effects believed to be thermal in origin, and that a satisfactory yield of matched silicon elements appears to be obtainable by the selection and resistance padding methods employed. Temperature control is used for the sake of constancy of the compression characteristic. Aging has been found negligible over a period of many months.

The actual compression characteristic afforded by these units is plotted

in Fig. 14 in terms of compressed signal amplitude vs. uncompressed signal amplitude, both in decibels.

Slicers. The slicer circuit shown in Fig. 15 resembles a conventional single-trip multivibrator in configuration, but functions somewhat differently because of the choice of parameters. In particular, capacitor C is made large enough so that the potential drop across it does not vary appreciably during normal operation, and plate resistor R_1 is given a small value such that the gain around the feedback path of the circuit is approximately unity when both triodes are in their active regions. Germanium varistors VR_1 and VR_2 maintain desired bias conditions regardless of the number of pulses or spaces in the input signal. Unlike the single-trip multivibrator, which when tripped remains so until the charge on C has had time to relax,

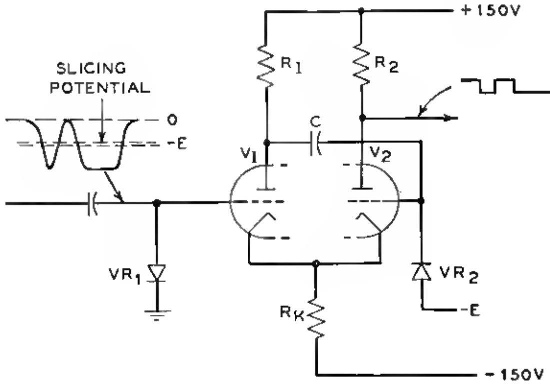


FIG. 15. Slicer circuit.

this circuit trips whenever the input signal falls through a narrow potential range near the value E, and trips back again when the input rises through the same potential range. A square wave of constant amplitude representing an accurate thin slice of the input is thus made available across resistor R_2. The circuit is capable of high speed, slicing pulses as narrow as 0.1 microsecond.

Time gating may readily be included among the functions of this circuit, through the addition of a triode having its plate and cathode connected directly to the plate and cathode, respectively, of tube V_1. The grid of this added triode is held normally at ground potential, and is pulsed in the negative direction by approximately rectangular gating pulses, having an amplitude of about $2E$. The total current passing through the common cathode resistor R_K is essentially constant, and if either or both of the paralleled tubes are conducting at a given instant this current is carried by one or shared by both of them. Hence the tripping action, involving transfer of

the cathode current to V_2, cannot occur as long as the grid of either V_1 or the added tube is positive with respect to the critical slicing potential. Tripping does occur whenever this limitation is removed, and thus the desired gating and slicing functions are performed concurrently.

In the experimental system this circuit is used to regenerate the PCM pulses at the common output of the two coders, and again at the input of each decoder to slice the received pulses and to sort out code groups of odd and even channels.

Coders. The method of binary coding used in this system was originally suggested by F. B. Llewellyn. It employs a novel electron beam tube. This device, pictured in Fig. 16, carries out simultaneously the two functions of quantizing and of coding. The tube is about $10\frac{1}{2}$ inches long and $2\frac{1}{4}$ inches in diameter. It has the 128 combinations of the 7-digit code laid out permanently as holes in an "aperture plate," and translates sample amplitudes from the form of beam deflections directly into PCM pulse symbols. Figure 17 shows one of these tubes in its socket on the rear of a coder panel, and above it a rectangular permalloy shield which covers the coding tube of the other coder serving the same 12-channel group.

As shown schematically in Fig. 18 the coder includes, in addition to the coding tube, a sampling and holding circuit which sorts out the odd (or even) channels from the input multiplex signal, push-pull amplifiers for vertical and horizontal beam deflections, and simple arrangements for blanking, focusing and centering. Within the tube are shown, from left to right, a conventional electron gun, vertical and horizontal deflection plates, a rectangular "collector" for secondary electrons, a "quantizing grid," the "aperture plate" and finally a "pulse plate." Figure 19 shows the target end of the tube, as viewed from a point near the gun. "Digit holes" in the aperture plate, laid out in accordance with the desired binary code, may be seen behind stretched parallel wires of the quantizing grid. One may count 64 narrow holes separated by equally narrow bars of metal in the left-hand vertical column, 32 holes in the next column, 16 in the next, and so on for seven columns. There are 129 grid wires uniformly spaced and accurately aligned so as to mask the upper and lower edges of every one of these holes when viewed from the geometric "point of origin" of the beam.

Stored audio samples from the sampling and holding circuit provide potential for the vertical deflection, with zero at the center, positive amplitudes in the upper half, and negative amplitudes in the lower half of the target area. A sawtooth sweep provides the horizontal deflection. The beam is blanked while deflection potentials are being changed to move it upward or downward from one sample amplitude to the next. When first restored, the beam strikes in the left-hand unperforated region of the aperture plate, and is then swept linearly across from left to right. Electrons which pass through the

Fig. 16. Coding tube. This new electron device transforms speech samples into pulse codes.

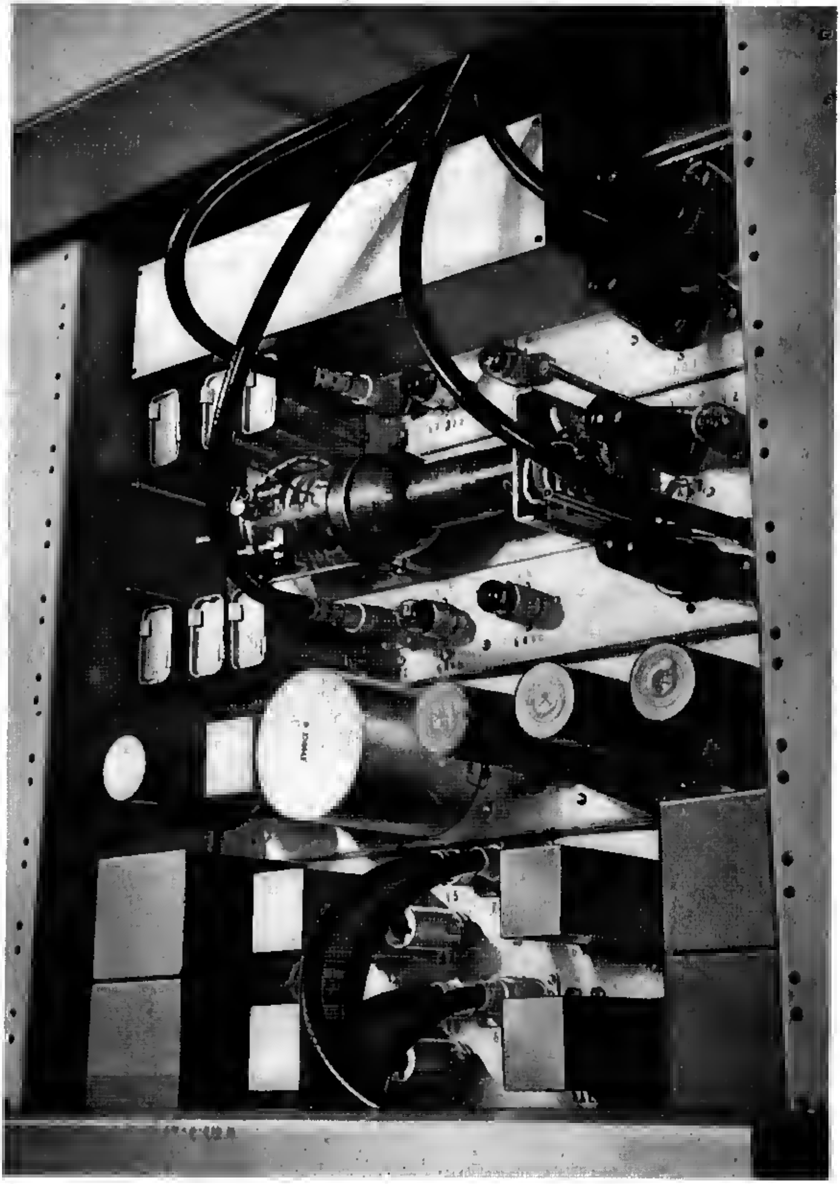


Fig. 17. Rear view of part of the collector (bottom panel), compressor, and two coders.

digit holes during the sweep are caught by the pulse plate, forming pulses which are amplified, gated and lengthened in the pulse regenerator to constitute the desired PCM signals. Retrace of the sweep occurs while the

beam is blanked, and is simultaneous with the application of the succeeding audio sample.

The wires of the quantizing grid are used to guide the beam so that it can illuminate only the particular row of apertures which correspond to the initial vertical deflection. Without this feature, erroneous codes would be produced when the beam straddled the edges of apertures or crossed from one amplitude level into another, as a result of electrode misalignment or possible slight drift in potential of an applied sample. The guiding action (basically proposed by W. A. Marrison and applied to the PCM coder at the suggestion of G. Hecht) is obtained by means of feedback from the quantizing grid to the vertical deflection amplifier.

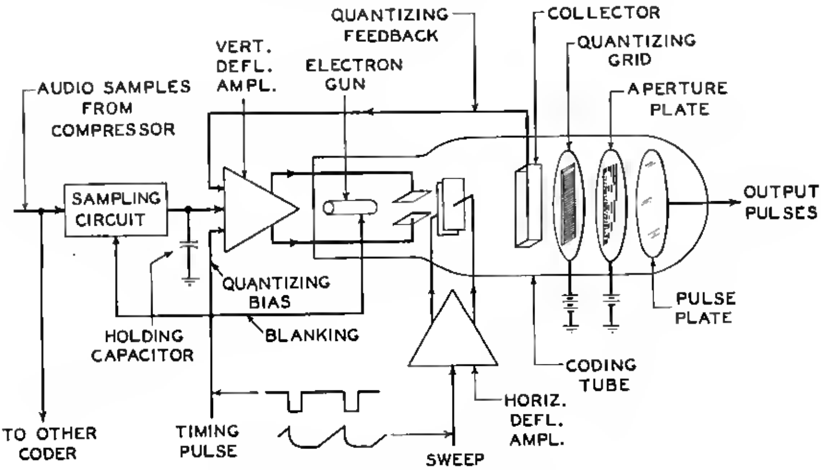


Fig. 18. Functional schematic of the coder.

The feedback signal is actually a current taken from the positively biased collector, which draws to it secondary electrons from the grid wires. The portion of the beam current striking the grid varies as a cyclic function of the vertical deflection. It follows that for some spot positions the value $\mu\beta$ in the feedback loop is positive, for others negative; hence with proper amplifier design there is a stable and an unstable region associated with each wire.

The spot can come to rest (vertically) only within one of the stable regions. In order to locate it consistently near the center of such a region, and thus gain equal margins against "hopping" upward or downward across a wire, a "quantizing bias" is introduced into the vertical deflection amplifier, along with the feedback and the signal samples. This bias is a current of opposite polarity from the unidirectional feedback current, and of magnitude equal

to the average between the two values of feedback current which exist when the beam falls (1) directly on a grid wire and (2) midway between two wires. One may regard this bias as pressing the beam upward against a wire, resisted by downward pressure associated with the feedback current. The latter

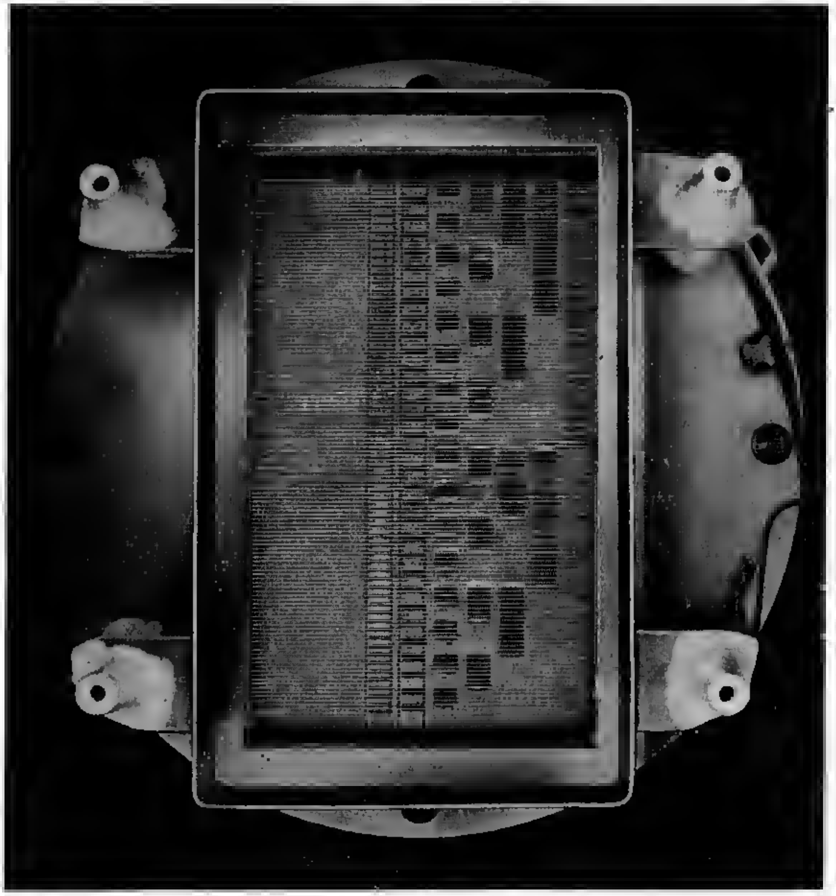


Fig. 19. Interior of the coder tube, viewed from the gun end.

increases as the beam approaches the wire, and equilibrium is reached when the beam is about half way between positions (1) and (2) mentioned above. The feedback current is actually intermittent, turned off and on by the blanking pulse, and careful analysis shows it necessary to make the bias intermittent also, with its wave fronts synchronous with those of the feedback current. This is readily accomplished by deriving the bias from the blanking signal itself.

With this arrangement the beam, suddenly turned on, moves either upward or downward from its initial "unquantized" position to the nearest position of stable equilibrium. Quantization is completed in less than a microsecond. Thereafter, as the beam is swept horizontally across the target area, it remains pressed upward against the lower surface of its guiding wire. Quantization is thus maintained until the end of the sweep, when blanking occurs. The margins against hopping across wires, while quanti-

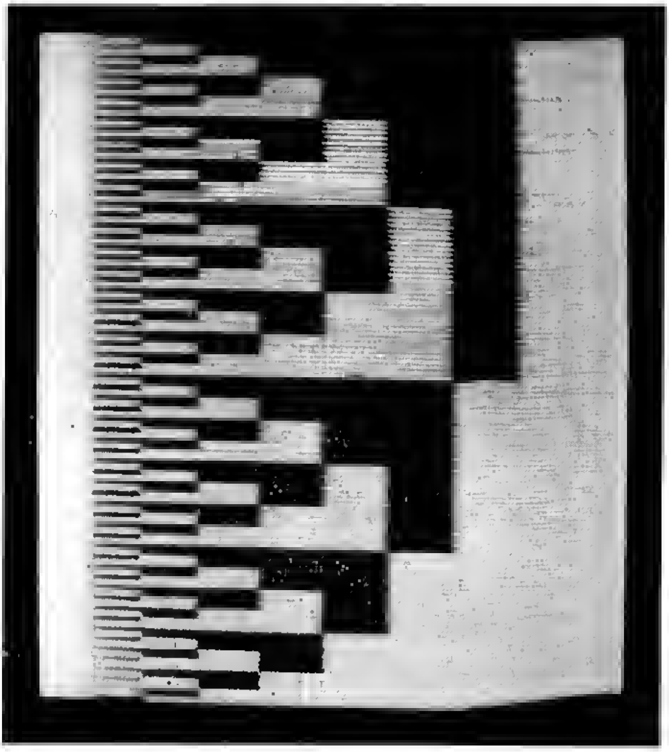


Fig. 20. "Television picture" of the aperture plate of a coder tube.

zation is in effect, are related to the amount of loop gain in the feedback path. In the present system feedback measuring about 20 decibels is provided, and will counteract changes which, without feedback, would move the beam two grid wires in either direction from the initial position.

This coding process, based on the electron beam coding tube and making use of a two-dimensional permanent layout of the code, is more straightforward than the various coding processes which depend upon counting or sequential comparison. Accordingly it leads to higher speeds and greater circuit simplicity.

Figure 20 illustrates the coding accuracy obtained. This photograph of

the screen of a test oscilloscope may be thought of as a television picture of the aperture plate of the coding-tube. To produce the pattern an audio-frequency sawtooth wave of full-load amplitude was applied to the sampling and holding circuit at the input of a coder. The resulting sample amplitudes, quantized by the coding tube and falling into all the possible 128 steps of the quantizing characteristic with uniform regularity, were used to energize the vertical deflection of the test oscilloscope. An ordinary synchro-

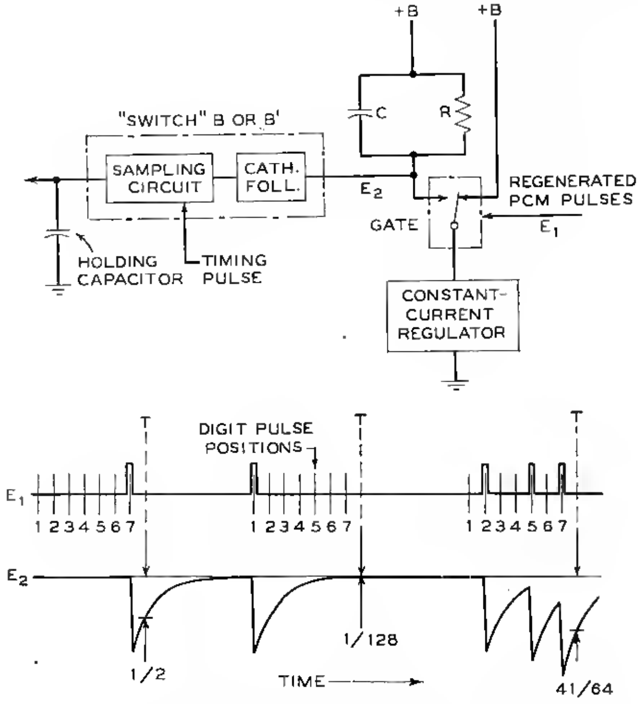


Fig. 21. Shannon decoding circuit and waveforms.

nized sweep provided the horizontal deflection. The square PCM pulses, delivered by the coder via the pulse regenerator, were applied to the intensity control.

Thus the code pulses corresponding to each quantized amplitude were made to appear as a row of blanks in a horizontal trace at the proper relative height. This pattern is very useful in studying coder performance.

Decoders. The decoding method is an impressively simple one originally proposed by C. E. Shannon. In its basic form it employs a pulsed resistance-capacitor circuit as illustrated in Fig. 21. Upon arrival of each pulse of the code, an identical increment of charge is placed upon the capacitor of the

device. The time constant $t = RC$ is such that, during any single pulse interval, whatever charge is on this capacitor decays precisely 50% in amplitude. It follows that the charge remaining at some chosen instant after the arrival of a complete code group consists of contributions of all its pulses, weighted in a binary manner. That is, if we define the contribution of a pulse in the final digit position as $\frac{1}{2}$, then contributions of $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, $\frac{1}{32}$, $\frac{1}{64}$ and $\frac{1}{128}$, respectively, are made by pulses in successively earlier positions. Any value from 0 to $\frac{127}{128}$, in steps of $\frac{1}{128}$, may thus be produced. Of course the digit holes in the aperture plate of the coder are laid out to make this straight-forward scheme workable. Since samples of low-level audio signals are coded near the center of the aperture plate, the corresponding decoded values lie in the neighborhood of $\frac{1}{2}$.

The basic Shannon decoder, then, comprises the resistance-capacitor circuit, means for supplying it with precisely controlled units of charge at precisely determined times, and a sampling and holding circuit (represented by switch B or B' in Fig. 10) to measure and store the decoded potential which is fleetingly present across the capacitor at a regularly recurring instant T, following the final pulse position. The scheme employed to inject the identical charges involves a regulated source of current and a gate to admit this current to the resistance-capacitor circuit under control of the regenerated PCM pulses. Two successive slicing operations and careful gating, as described earlier, make these pulses more than adequately uniform.

The wave form sketches of Fig. 21 show three typical decoding cycles. In the first, a single pulse in digit position 7 produces a decoded amplitude of $\frac{1}{2}$; in the second, a pulse in position 1 gives $\frac{1}{128}$; and in the tbird, pulses at 2, 5 and 7 provide a decoded value of $\frac{41}{64}$. It may readily be verified that the provision of an idle channel interval between operations (following from the alternate use of two decoders) allows the residue of one decoding operation to decay to a negligible value (never larger than $\frac{1}{128}$ of a single step height or "quantum") by the time of the next consecutive sampling. Experimentally, interchannel crosstalk from this source is virtually nonexistent.

In the foregoing it has been emphasized that precise timing is required for this basic Shannon decoder. Although the necessary accuracy was actually obtained without great difficulty in early tests, a modification has also been introduced which eases the requirements to a very marked extent. This scheme, devised by A. J. Rack, employs a damped resonant circuit in conjunction with the resistance-capacitance elements in a manner illustrated by Fig. 22. The natural frequency of the resonant circuit L, C_2, R_2 is made equal to the PCM pulse rate, and the time constant of the damped oscillation ($t=2R_2C_2$) is matched to that of the circuit R_1, C_1. The same charging

pulses pass through both sections of the circuit; hence by proper choice of C_1 and C_2 the amplitudes of the damped sine wave and the exponential may be proportioned so that the rate of change of their combined potential becomes zero at successive points one pulse period apart. In fact it has been found possible to make both the first and the second time derivatives of potential equal to zero simultaneously at such points.

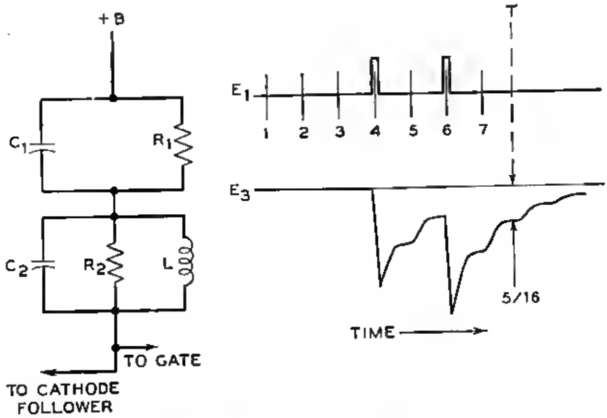


Fig. 22. Shannon-Rack decoder.

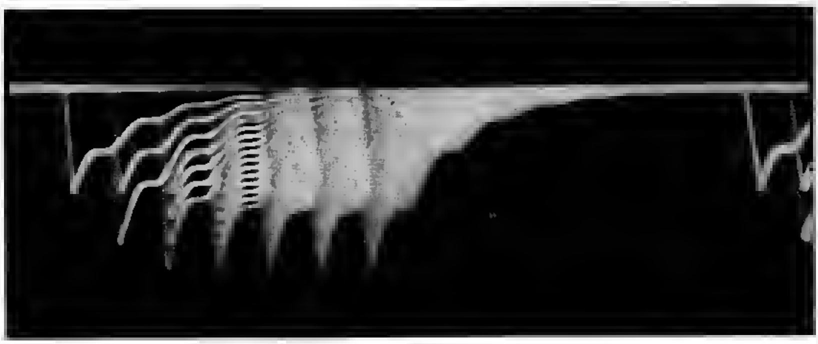


Fig. 23. Output of Shannon-Rack decoder for a signal giving 100% modulation.

By this modification, not only the times of application of the charges but the time of sampling is made much less critical. Of course this presumes that the sampling circuit is designed to complete its operation near the center of a level region. In Fig. 22, the voltage transient due to a typical pair of pulses is sketched, and Fig. 23 shows an oscilloscope screen, on which the waveforms delivered by the cathode follower of one of the Shannon-Rack decoders of the system are superimposed for a full set of 128 possible code combinations in sequence.

V. Performance

In general the behavior of the system has shown promise for toll plant application. Among other things the stability realized in the adjustments of the coder and decoder, and the apparent absence of aging or other drift in the compressor and expandor have been gratifying. A daily check of the focus and centering in the coders and of the time constants in the decoders appears adequate to keep them in optimum adjustment. The synchronizing

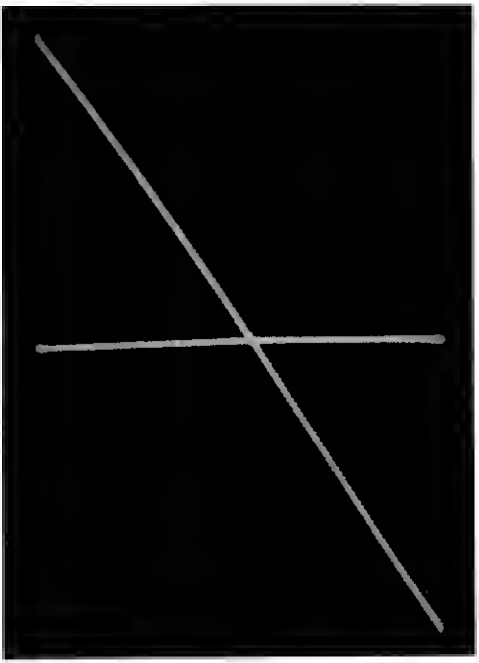


Fig. 24. Output of decoder (vertical) vs. input to coder (horizontal).

gear has been found equally easy to maintain. On the other side of the ledger, occasional breakdowns have pointed up the need for alarm and automatic replacement facilities in any version of the system which might be developed for commercial service.

A few measured characteristics are given in this section in addition to the compression and audio gain-frequency characteristics already shown (Figs. 14 and 11, respectively).

Input-Output Characteristics. The diagonal trace in the oscilloscope pattern of Fig. 24 shows the relationship between the input of a coder (horizontal deflection) and the output of the corresponding decoder (vertical). For this pattern a full-load audio signal was applied to the input of the odd

coder only (without passing through the compressor), and the output was taken directly from the common output of the two decoders. Thus the six odd channels took turns transmitting the signal while the six even channels produced the horizontal center-line. Uniform quantizing steps may be seen along most of the trace, but are obscured near the ends by defocusing of the test oscilloscope.

A similar pattern, obtained with the compressor and expandor included in the transmission path, is shown in Fig. 25. Here tapered steps may be discerned, as well as a small amount of non-linearity due to residual imper-

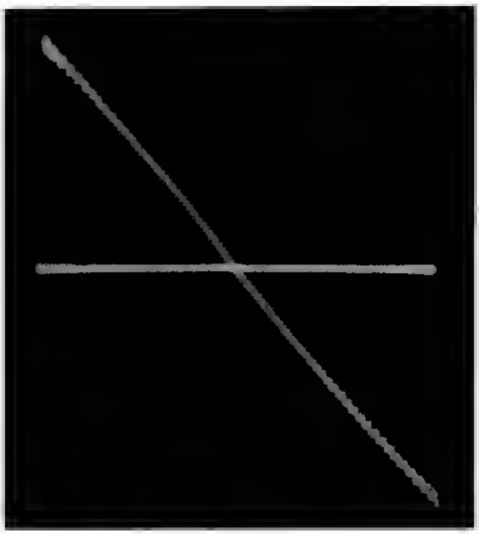


Fig. 25. Output of expandor (vertical) vs. input to compressor (horizontal).

fections in the companding. The effect of the channel peak choppers is not included in this pattern.

Two measured overall input-output characteristics appear in Fig. 26, for the case of a typical single channel and for five channels patched in tandem through 17-decibel pads on a 4-wire basis. The latter simulates a possible extreme case of a long circuit in which for some reason it is desired to decode to audio at four junction points as well as at the final terminal. It should not be confused with a series of spans between which the PCM pulses are amplified or regenerated without decoding. In the latter case, of course, the overall audio characteristics are independent of the number of spans.

Audio Noise. Quantizing was found to be the only significant source of noise in the received audio signals. Noise levels measured in the absence of speech are shown in Fig. 27. The measurements are given for various num-

bers of channels from one to ten, connected in tandem as described in the preceding paragraph. Two scales of ordinates are shown in this figure. On

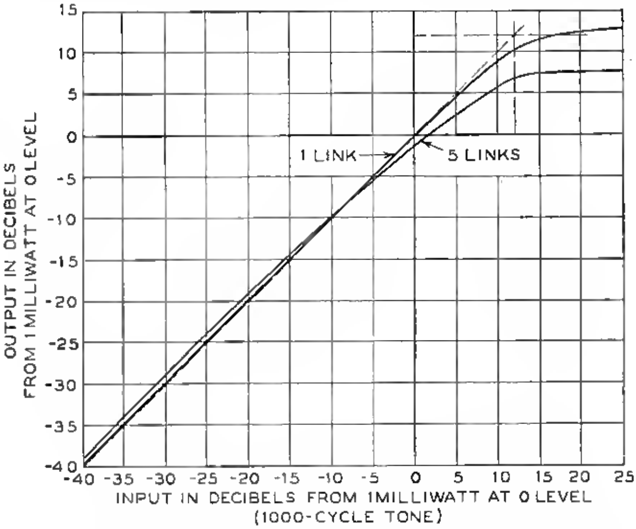


Fig. 26. Input-output characteristics of PCM channels.

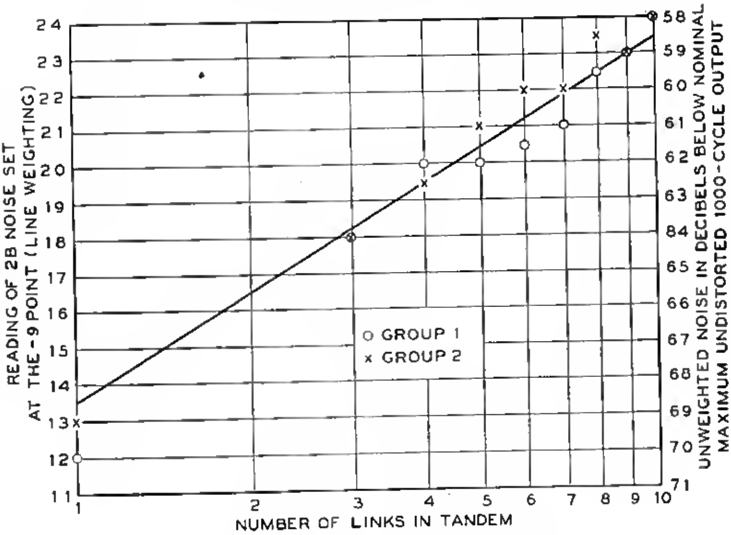


Fig. 27. Noise measurements on idle PCM channels.

the left is a reference scale of weighted noise employed in the Western Electric 2B Noise Set; on the right, a scale which relates the corresponding unweighted root-mean-square noise to the nominal maximum undistorted

single-frequency output of the system (taken to be 9 decibels above a milliwatt at zero level). This output corresponds to an input just reaching the peak limiters. Thus in a single link the idle circuit noise is down about 68 decibels from the full-load sine wave. For five links in tandem, the readings are at least 8 decibels below the accepted limit (29 decibels on the left-hand scale) for noise at the end of a 4000-mile circuit. Quantizing noise is found to increase approximately three decibels for each doubling of the number of links, as is generally the case with other forms of random noise.

For a single channel, or a small number of channels in tandem, the idle circuit noise varies considerably with the vertical centering of the coding tube. This may be understood by noting that if the zero-signal operating point effectively rests near the center of a "tread" on the quantizing staircase, (Fig. 4) a small amount of power hum or other disturbance may simply move it back and forth on the same tread, in which case quantizing noise is entirely absent. On the other hand, if the operating point is near a "riser" the small disturbances may cause it to joggle from one tread to another, producing noise. The measurements given in Fig. 27 were obtained with a very quiet input circuit and with the centering adjusted for maximum noise. The "joggling" was produced largely by residual power hum.

A-B tests to compare PCM transmission over a single link with direct transmission over a noise-free circuit of the same audio band were carried out using a wide range of talker volumes. In such tests only a few experienced observers were able to pick the PCM path consistently. When the PCM circuit included five tandem links most observers could tell the difference, but all judged the quality to be more than satisfactory for toll service.

Crosstalk. It has been pointed out earlier in this paper that interchannel crosstalk in a PCM system can occur only in the terminal equipment. Considerable care was exercised, particularly in the design of the time-division parts of the system, to hold individual sources of crosstalk to 70 decibels or better. As a result, measurements using a single-frequency test tone and a current analyzer have shown the overall crosstalk from any one channel to any other to be down 66 decibels in the worst cases.

Very severe tests have also been made in which a loud talker was connected to ten channels of a group simultaneously and crosstalk into either of the two remaining channels (one odd, one even) was listened to, and measured with the 2-B Noise Set. In such tests unintelligible crosstalk could be detected, which seemed to consist of changes in the quality of the quantizing noise occurring at a syllabic rate. The 2-B readings averaged about a decibel above the quantizing noise of a single idle channel with occasional peaks reaching the 17-decibel point on the reference scale.

In tests involving 24 talkers in 12 simultaneous conversations, crosstalk was practically undetectable.

Radio Interference and Noise. To obtain experimental confirmation of the expected tolerance to high interference levels in the radio path, the output of an oscillator, tunable through the band near 65 megacycles, was superimposed upon the received intermediate-frequency signal at the input to the group selection filters. With this controllable interference tuned near the center of the Group 1 filter, and its amplitude set 6 decibels below the peak amplitude of the (noise-free) pulses, errors were so plentiful that the demodulator did not remain synchronized. Proper framing was restored when the amplitude difference was increased to 7 decibels, but enough decoding errors remained to give intolerable audio noise. At 8 decibels only an occasional crackle of noise was observed, and at 9 decibels reception was perfectly normal. Similar tests of Group 2 gave the same results except, of course, that synchronization was not affected. The fact that noise-free transmission was not maintained quite up to the ideal 6-decibel point is due principally to the width (0.4 microsecond) of the gate which is applied to the rounded PCM pulses entering the receiving equipment. If necessary the ideal could undoubtedly be approached more closely by reducing this width, thus admitting only the extreme peaks and troughs of the signal.

The effects of actual fluctuation noise were studied by sending the PCM pulses over the radio path at reduced level. The boundary between good and bad transmission was not so sharp as with the continuous-wave interference, as should be expected because of the random nature of the noise. Flawless reception occurred when the root-mean-square signal at the peak of a pulse was greater by 18 decibels than the root-mean-square noise, both measurements being made at the output of a group selection filter.

VI. Acknowledgment

The system described is the result of the co-ordinated efforts of many people. In particular, the writers wish to express their appreciation to Mr. R. W. Sears for his development of the coding tube, and to acknowledge the contributions of certain others who have been actively concerned with various phases of the project. These include Messrs. R. L. Carbrey, A. E. Johanson, J. M. Manley, G. W. Pentico, A. J. Rack, P. A. Reiling, L. R. Wrathall, and the mechanics and wiremen who have kept to their usual high standards throughout countless circuit changes.

The project was under the overall direction of Mr. C. B. Feldman.